

GRAMMIS – EIN PROJEKT STELLT SICH VOR

von Eva Breindl, Roman Schneider, Bruno Strecker

GRAMMIS – kurz für: Grammatisches Informationssystem – ist ein Forschungsvorhaben der Abteilung Grammatik des Instituts für Deutsche Sprache in Mannheim, das zum Ziel hat, Grammatik unter Nutzung neuer Medien und Techniken multimedial und als Hypertext vernetzt aufzubereiten.

Was und wozu?

Wozu dieser Aufwand? Immerhin wurde in derselben Institution erst vor wenigen Jahren eine Grammatik in drei dicken Bänden¹ abgeschlossen. Das neue Projekt tritt zu diesem Werk in keiner Weise in Konkurrenz, im Gegenteil: Es baut in weiten Teilen auf ihm auf und wäre ohne dessen Leistung sicher gar nicht möglich geworden. Tatsächlich verdankt sich die Projektidee selbst Erfahrungen, die sich bei der Arbeit an dieser Grammatik in traditioneller Buchform ergaben und die multimediale Hypertexte in vieler Hinsicht als ideale Form für die Präsentation grammatischer Informationen erscheinen lassen.

Wer je damit befasst war, eine Grammatik oder auch nur einen größeren Ausschnitt einer solchen zu schreiben, kennt, jenseits aller Sachfragen, diese Schwierigkeiten:

- Unter welchen Aspekten sollen die Informationen zusammengestellt werden?
- Wie soll man den Text aufbauen?
- Für wen schreibt man überhaupt?

Was formal gleicher Art ist, kann aus funktionaler Sicht verschieden sein. Was von seiner kommunikativen Funktion her zusammengehört, ist oft genug formal so heterogen, dass die Tradition geradezu verbietet, den Aufbau

einer Grammatik daran zu orientieren. Zusammen führen die verschiedenen Klassifikationsaspekte zu echter Doppelbindung. Wie immer man sich entscheidet, wichtige Gesichtspunkte bleiben unbeachtet. Behilft man sich mit ständigen Querverweisen, belastet dies die ohnehin schwierige Nutzung des Textes in unzumutbarer Weise. In einem Buch ist die beste Lösung noch, beide Betrachtungsweisen in eigenen Angängen nebeneinander zu stellen.

Grammatiken sind – wie immer man sie gestaltet und wie immer sie theoretisch orientiert sein mögen – so komplex, dass ihr Textaufbau selbst zum Problem wird, weil benötigte Informationen nur von Kennern mit vertretbarem Aufwand aufzuspüren sind. Sach- und Wortverzeichnisse helfen hier nur beschränkt. Dass Laien, die sprachliche Probleme lösen müssen, mit herkömmlichen Grammatiken in der Regel überfordert sind, ist auch eine Erfahrung der Sprachberatungsdienste. Mehr Transparenz und damit Verständlichkeit ist nur durch Reduktion der angebotenen Information zu erreichen, doch dem steht entgegen, dass in Grammatiken gerade solche Informationen besonders gefragt sind, die Spezialfälle betreffen.

Nicht weniger schwierig ist die adressatengemäße Darstellung. Das Spektrum möglicher Adressaten reicht vom ungeübten, gelegentlichen Schreiber bis zum Fachkollegen, der sich punktuell für eine bestimmte grammatische Erscheinung interessiert oder gar nur sehen will, wie andere behandeln, was er selbst untersucht. Da es so gut wie ausgeschlossen ist, so unterschiedlichen Interessen und Wissensvoraussetzungen mit einem Werk zu genügen, hat sich der Markt für Grammatiken in Segmente aufgespalten, die bestimmte Benutzergruppen gezielt zu bedienen suchen. Das ist grundsätzlich nicht schlecht, doch ist es selten wirklich gelungen.

Das Aufkommen von Hypertextsystemen hat hier eine völlig neue Lage geschaffen, denn sie erlauben, im Rahmen eines kohärenten Systems Informationen unterzubringen, die gleichermaßen verschiedene Perspektiven, verschiedene Interessen und verschiedene Wissensvoraussetzungen in Rechnung stellen können. Allein schon deshalb sind Hypertextsysteme ideale Werkzeuge für die Grammatikschreibung. Hinzu kommt, dass sie Möglichkeiten eröffnen, die man auf Grund der natürlichen Beschränkungen von Druckerzeugnissen schon ganz aus den Augen verloren hatte:

- In Hypertexten können Tondokumente eingebunden werden, was in Anbetracht des Primats des Mündlichen vor dem Schriftlichen von elementarer Bedeutung ist. Insbesondere können etwa Betonungsverhältnisse nicht nur beschrieben, sondern auch zu Gehör gebracht werden.
- Was als Vorgang beschrieben werden soll, kann mit Hilfe animierter Bilder anschaulich gemacht werden, so etwa das Verhältnis von Hauptsatz- und Nebensatzstellung.
- Die Seiten eines Hypertextes sind nicht statisch wie die eines gedruckten Textes. Sie können an Nutzerprofile angepasst werden und dynamisch auf Nutzereingaben reagieren, etwa indem sie Informationen aus Datenbanken abrufen.

GRAMMIS nutzt das ganze Spektrum dieser Möglichkeiten: Bei zahlreichen Beispielen kann per Mausklick das entsprechende Tondokument aufgerufen werden. Vorgangsbeschreibungen werden mit Hilfe sog. JAVA-Applets mit bewegten Bildern illustriert. Formulare mit vom Nutzer auswählbaren oder ausfüllbaren Optionsfeldern erlauben interaktive Recherchen, insbesondere auch Recherchen in der Bibliografie zur deutschen Grammatik.² Ein Beispiel hierfür zeigt Abb. 1.

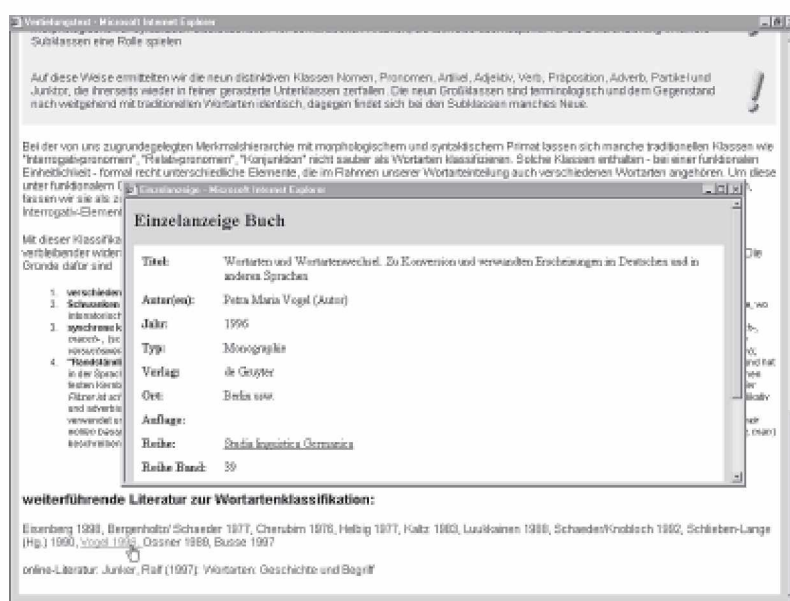


Abb. 1

Die technische Konzeption

Die grundlegende Entscheidung, GRAMMIS als hypermediales Informationssystem im World Wide Web (WWW) zu präsentieren, hatte von Anfang an Auswirkungen auf die technische Konzeption und Realisierung des Projekts. Von wesentlicher Bedeutung sind in diesem Zusammenhang die medienspezifische Aufbereitung und Darstellung der Inhalte, die explizit nicht-linear miteinander verknüpft und – wo angebracht – um multimediale Komponenten ergänzt werden sollen, sowie der Anspruch, einmal erstellte Inhalte bei Bedarf auch für andere Publikationsformen nutzen zu können.

Die zentrale Idee kann unter dem Schlagwort »Infopool« zusammengefasst werden. Ganz gleich, ob das WWW oder ein potenzielles Nachfolgemedium in einigen Jahren auf ganz anderen als den heutigen Standards aufbaut oder ob sich die aktuellen Techniken – beispielsweise bei der Auszeichnung von Texten – im globalen Kontext behaupten können: Einmal geschriebene GRAMMIS-Texte müssen weiterhin maschinell verarbeitbar und damit nutzbar sein. Außerdem sollen die Autoren die Möglichkeit haben, ihre Arbeiten ohne allzu großen Aufwand auch in anderen Medien (Printmedien, Computer-CD etc.) publizieren zu können.

Aus solchen Gründen liegt es nahe, ein Informationssystem wie GRAMMIS unter Beachtung möglichst offener, weltweit akzeptierter Standards sowie unter Verwendung leistungsfähiger und robuster Datenhaltungsmechanismen zu realisieren. Im Einzelnen bedeutet dies:

- **Datenformat:** Elementar ist hier die Idee des »logical markup« als Gegenstück zum »visual markup«. Ziel ist die explizite Trennung von Inhalt und äußerer Erscheinungsform eines Dokuments. Umgesetzt wird dies in Form einer für GRAMMIS maßgeschneiderten Auszeichnungssprache auf der Basis von XML (eXtensible Markup Language), dem vom World Wide Web Consortium (<http://www.w3c.org>) definierten Standard zur Beschreibung unterschiedlichster Textsorten. Den Autoren wird damit ein jederzeit erweiterbares Hilfsmittel zur eindeutigen Markierung von Überschriften, Absätzen, Gliederungslisten oder Fußnoten an die Hand gegeben.

Anders als Texte, die unter Verwendung layoutorientierter Auszeichnungssprachen wie HTML oder WYSIWYG-Textverarbeitungen wie Winword erstellt wurden, lassen sich XML-Dokumente einfach überarbeiten, flexibel darstellen und für unterschiedliche Ausgabemedien formatieren sowie maschinell analysieren und auswerten. Insbesondere der letzte Punkt ist für die Entwicklung ambitionierter Recherchemöglichkeiten interessant. GRAMMIS-Dokumente werden nach Vorgaben strukturiert, die in einer speziellen »Document

Type Definition« (DTD) – einer kontextfreien Grammatik – festgelegt sind. Für die medien- und kontextspezifische Darstellung sowie die Transformation in andere Formate werden flexibel anpassbare Formatierungsregeln („XSL-Stylesheets“) eingesetzt. Nicht-textuelle Inhalte (z.B. Animationen) werden mit Hilfe der plattformunabhängigen Programmiersprache Java erstellt und in die XML-Dokumente eingebettet.

- **Datenhaltung:** Wer in größerem Maßstab mit und für den Computer publiziert, weiß um die Schwierigkeiten und Gefahren des elektronischen Mediums. Was tun, wenn die Menge der zu verwaltenden Dokumente ins Unüberschaubare wächst oder einzelne Dateien auf Grund von Datenträgerfehlern plötzlich nicht mehr verwendbar sind? In Mehrbenutzer-Umgebungen kommt noch der Aspekt verteilter Bearbeitung der Ressourcen hinzu: Welche Dokumente sollen wann und für wen erreichbar sein, wie wird den Gefahren unbeabsichtigten gleichzeitigen Zugriffs vorgebeugt und dergleichen mehr.

Die Lösung liegt hier in der Verwendung eines leistungsfähigen Datenbankmanagementsystems (DBMS), das einerseits »im Hintergrund« für Datensicherheit, Zugriffskontrolle und Datenkonsistenz sorgt und andererseits mächtige Werkzeuge für die Inhaltsanalyse und Benutzerrecherche anbietet. Im Rahmen des Projekts GRAMMIS wird das objektrelationale DBMS Oracle eingesetzt.

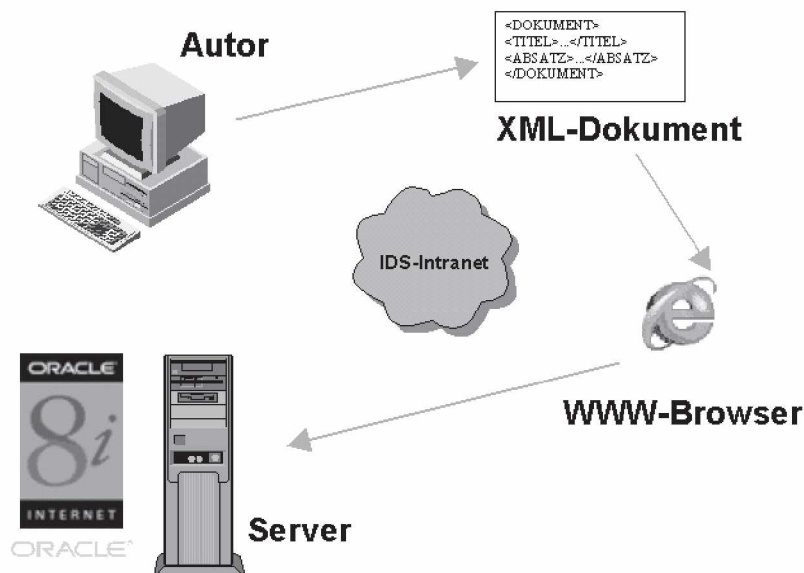


Abb. 2 verdeutlicht die technische Konzeption von GRAMMIS sowie speziell die Arbeitsweise der Autoren: Mit Hilfe eines Editionswerkzeugs, das sie bei der Einhaltung der Strukturregeln unterstützt (»Validierung«), erstellen sie am lokalen PC Hyperdokumente und kümmern sich dabei allein um deren logische Strukturierung. Das spätere Layout muss zu diesem Zeitpunkt noch nicht festgelegt werden. Mittels eines Web-Browsers werden die lokal er-

stellten Dateien in die zentrale GRAMMIS-Datenbank »eingespeichert«. Dort werden anschließend Prozeduren angestoßen, welche die gespeicherten Hyperdokumente miteinander verknüpfen, indizieren und für spätere Recherchen vorbereiten. Wichtig ist dabei, dass die Inhalte für ein potenziell heterogenes Publikum verfügbar sein sollen, also automatisch auf Grund individueller Nutzerprofile aufbereitbar und präsentierbar sein müssen.

Der Zugriff auf GRAMMIS-Inhalte setzt auf Anwenderseite lediglich einen aktuellen WWW-Browser voraus. Der Anwender identifiziert sich über ein persönliches Passwort und erhält Zugang zu den textuellen Komponenten und zu den ergänzenden Datensammlungen, die mittels Hyperlinks untereinander verbunden sind. Auch wenn die originäre Art der Informationssuche in Hypertextnetzen das Stöbern (»Browsing«) ist, muss ein leistungsfähiges Informationssystem gezielte Suchanfragen unterstützen. In diesem Zusammenhang werden weitere Vorteile der Speicherung in einem DBMS deutlich: Mit Hilfe der Datenbanksprache SQL können komplexe Abfragen unter einer nutzerfreundlichen und einfach bedienbaren Oberfläche angeboten werden, so etwa:

»Zeige mir alle Dokumente, die von Autor X zwischen Mai 1999 und Februar 2000 erstellt wurden, das Thema Y behandeln und das Wort Z beinhalten. Bereite diese Dokumente entsprechend meines Nutzerprofils auf und zeige mir zusätzlich alle thematisch relevante Literatur sowie in diesem Kontext interessante Einträge aus dem Wörterbuch.«

Neben themenbezogenen Recherchen, die auf dokumentspezifischen Meta-Informationen (etwa inhaltlich klassifizierende Schlagwörter) sowie einer automatischen Analyse der Dokumentenstruktur und der Position im Hypertext-Netzwerk basieren, wird eine Reihe weiterer Optionen für die Informationssuche geboten. Dazu zählen eine exakte Wort- und Phrasensuche unter Verwendung von Platzhalter-Zeichen und Booleschen Operatoren, eine Fuzzy-Suche nach ähnlich geschriebenen oder klingenden Ausdrücken (Letzteres auf Basis einer phonetischen Analyse durch einen »Soundex«-Algorithmus) sowie eine automatische Bewertung der Suchergebnisse (»Scoring«).

Was GRAMMIS bieten wird

Einmal fertig gestellt, wird das System diese Komponenten haben:

Grammatisches Grundwissen vermittelt eine gleichnamige Komponente in gewohnt systematischer Form, jedoch multimedial aufbereitet und in drei Informationsebenen geschichtet: Damit man sich leicht und schnell orientieren kann, wird komplexeren Informationseinheiten ein

Kompakttext vorangestellt, von dem aus man per Mausklick zu **detaillierten Darstellungen** gelangt. Wer weiter einsteigen will, findet, wo dies angebracht erscheint, Übergänge zu **vertieften Informationen**. Wie man sich dies vorzustellen hat, zeigt dieses Beispiel aus der Einheit »Wortarten« (Abb. 3):

Auf der Kompaktebene präsentiert sich eine unkommentierte anklickbare Liste der Wortarten des Deutschen, die zu Überblicks-Darstellungen für einzelne Wortarten führt.

Auf der Detailebene werden die Klassifikationskriterien genannt, wird das kreuzklassifikatorische Verfahren vorgestellt und die vorgenommene Einteilung begründet. Verweise führen zu den Detail-Darstellungen für die einzelnen Wortarten.

Dort finden sich Verweise auf die Vertiefungsebene. Einer davon führt zu einem graphischen Überblick über die Wortarten und deren wesentliche differenzierende Merkmale. Diese Information setzt durch die verkürzte Merkmalschreibweise einiges Insiderwissen voraus.

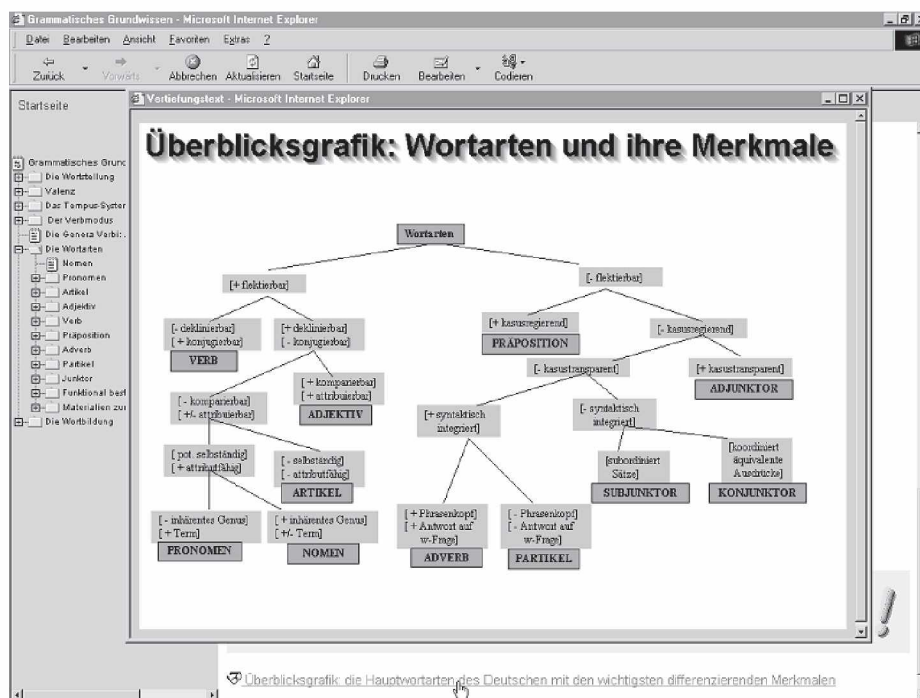


Abb. 3

Wer Probleme hat, seine Fragen auf einen terminologischen Nenner zu bringen, kann über eine Liste Hilfe finden, in der **häufig gestellte Fragen** zur deutschen Grammatik zusammengestellt und unter Verwendung von reichem Beispielmateriale möglichst leicht verständlich beantwortet werden. Hyperlinks leiten weiter zu einschlägigen Einheiten des Grammatischen Grundwissens.

Wo sich eine Anfrage an einem Wort festmachen lässt (etwa: »Kann man sagen *weil ich weiß das nicht?*«), kann das **Grammatische Wörterbuch** weiterhelfen, das Angaben zu grammatisch relevanten Wirkungen von Wörtern wie *weil* auf ihr Umfeld in Sätzen enthält.

Wer sich über den Einzelfall hinaus für knifflige grammatische Fragen interessiert, ohne gleich aufwändige Grammatikstudien betreiben zu wollen, kann zu den **Hauptschwierigkeiten der deutschen Grammatik** gehen, die sich auf Fragen konzentrieren, die auch für fortgeschrittene Deutschlerner von Interesse sind.

Eine **Grammatikexpertin** – ein terminologisches Glossar – bietet kompakte Erklärungen grammatischer Fachausdrücke und Übergänge zu detaillierteren Beschreibungen und einschlägiger Fachliteratur. Über so genannte Hotwords – das sind speziell markierte Ausdrücke – kann man aus GRAMMIS-Texten aller Art heraus per Mausklick direkt auf dieses Glossar zugreifen, ohne den aktuellen Text zu verlassen (siehe Abb. 4).

Umfangreiche **Bibliographien zur deutschen Grammatik**, allgemein und zu speziellen Themenbereichen, runden das System ab.

Wann und wo?

Wann ist mit GRAMMIS zu rechnen? Ein Anfang ist gemacht. GRAMMIS ist seit Anfang 2000 unter der Adresse <http://www.ids-mannheim.de/grammis> im WWW zugänglich, freilich nicht in vollem Umfang. Systeme der Komplexität von GRAMMIS lassen sich nicht im Handumdrehen erstellen, auch wenn man modernste Rechner zur Verfügung hat. Was später schnell und mühelos zu erledigen sein soll, muss zunächst einmal mühsam und mit viel Handarbeit eingegeben werden. Derzeit stehen deshalb erst diese Teilkomponenten der Komponente **Grammatisches Grundwissen** zur Verfügung, deren Entwick-

lung Vorrang hat, weil sie den informatischen Kern des gesamten Systems bildet³:

Das Tempussystem des Deutschen

Vorgestellt wird das System der Tempusformen deutscher Verben in einer Weise, die mit der traditionellen Einteilung nicht viel mehr gemeinsam hat als die Anzahl ihrer Stufen: Präsens, Präteritum, Präsensperfekt, Präteritumperfekt, Futur, Futurperfekt. An die Stelle eines simplen Früher, Jetzt und Künftig tritt eine Betrachtung, die Tempusformen im Hinblick darauf interpretiert, aus welcher zeitlichen Warte sich Sachverhalte präsentieren, wenn eine bestimmte Tempusform gewählt wurde. Die zum Teil recht komplizierten Beziehungen zwischen Sprechzeit, Betrachtzeit und

Ereigniszeit werden mit Hilfe animierter Graphiken anschaulich gemacht.

Der Verbmodus

Beschrieben und semantisch interpretiert werden der Halbmodus Imperativ – halb, weil in den Dimensionen Tempus und Person nicht differenziert – und die beiden Vollmodi Indikativ und Konjunktiv, wobei der Schwerpunkt auf der Darstellung des Konjunktivs liegt. In einer Breite, die sonst nur in speziellen Untersuchungen zu finden ist, werden die Verwendungsbedingungen der Modi betrachtet. Besondere Beachtung findet dabei die Rolle von Indikativ und Konjunktiv bei der indirekten Redewiedergabe. Authentische Tonbeispiele belegen den Gebrauch von Konjunktiven in verschiedenen Verwendungskontexten.

Die Wortarten des Deutschen

In dieser Einheit werden Wörter auf der Grundlage eines Kriterienbündels aus morphologischen, syntaktischen und semantisch-funktionalen Eigenschaften klassifiziert und beschrieben. Da in der Regel auch Nutzer mit geringeren Vorkenntnissen mit irgendeinem Wortartensystem Bekanntschaft gemacht haben, kommt dieser Einheit eine wichtige Rolle als »Einstieg« zu. Die Anbindung an andere Klassifikationen und Termini ist sicherzustellen: Zum einen muss, auch wer etwa nach »Substantiv« oder »Hauptwort« sucht, Informationen zur Wortart der Nomina finden, zum anderen sollte ausgehend von den einzelnen Wortarten schrittweise der hierarchische Aufbau von Sätzen über Wortgruppen (Phrasen) und syntaktische Funktionen nachzuvollziehen sein.

Die hypertextuelle Segmentierung der Wortarten ergibt sich aus dem Gegenstand: Jede Wortklasse und Subklasse wird als eigene Informationseinheit behandelt, was zum Teil zu längeren Einheiten führt. Diese werden dann durch textinterne Sprungmarken über eine einheitliche Gliederung nach den morphologischen, syntaktischen und semantisch-funktionalen Merkmalen erschlossen.

Die Wortbildung im Deutschen

In der **Grammatik der deutschen Sprache** findet sich kein Überblick über die Wortbildung des Deutschen. Er wird für GRAMMIS neu entwickelt. Weitgehend

abgeschlossen sind die definitorischen Grundlagenkapitel (»Die Einheiten der Wortbildung«, »Die Wortbildungsarten«), die bereits durch zahlreiche

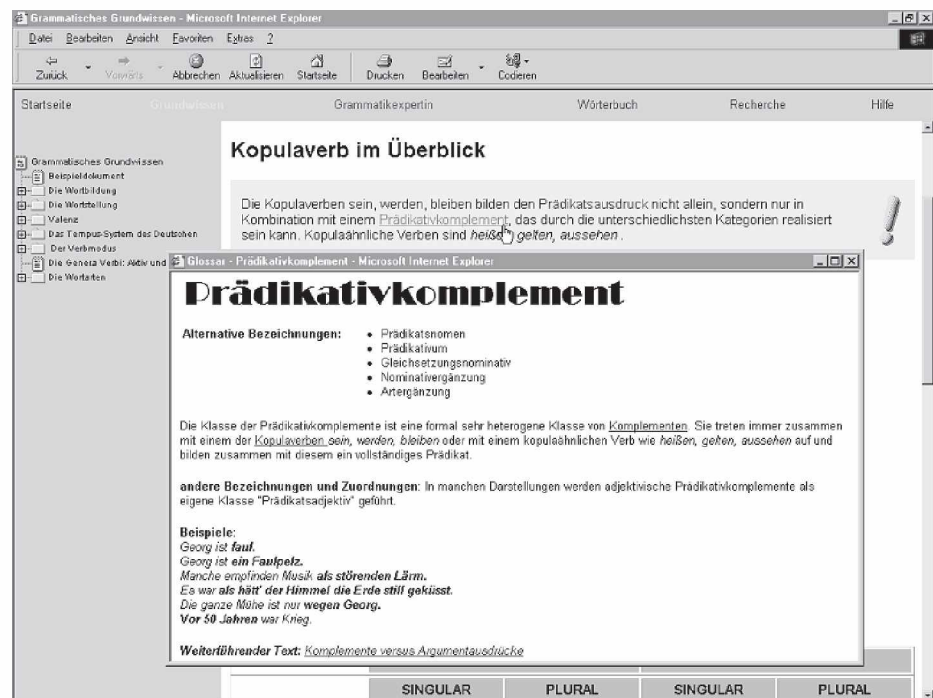


Abb. 4

Glossareinträge ergänzt werden. Es ist vorgesehen, in dieser Komponente auch die semantischen Muster vorzustellen. Im Grammatischen Wörterbuch sollen darüber hinaus alle Wortbildungsaffixe des Deutschen kurz beschrieben werden.

Die Wortstellung im Deutschen

Bei der Darstellung der Stellungsregularitäten von Ausdruckseinheiten in Sätzen zeigt sich der Mehrwert eines multimedialen Hypertextes in besonderem Maß. Wo immer Änderungen in der Linearstruktur zu beschreiben sind, wo Komponenten ihren Ort wechseln, kann dies durch animierte Graphiken wiedergegeben werden. Der beachtliche Aufwand beim Erstellen solcher Animationen zwingt freilich zu sparsamem Umgang mit diesen Möglichkeiten. In einer Untereinheit werden die Wortstellungsphänomene zusammengefasst und mit Tonbeispielen belegt, die in besonderer Weise sensitiv für Intonation und Betonung sind. Solche Tonbeispiele vermitteln einen holistischen und realistischeren Eindruck der akustischen Determinanten der Wortstellung als dies mit graphischen Hilfsmitteln wie Unterstreichungen für Betonung oder Pfeile für die Tonhöhenbewegung jemals möglich wäre.

In Arbeit und in Teilen bereits nutzbar sind auch die Komponenten Grammatikexpertin und Grammatisches Wörterbuch. Ihre Struktur liegt fest, ebenso die Formulare für Abfragen. Das Glossar erklärt derzeit rund 150 Fachbegriffe, im Wörterbuch fehlen noch die Massen von Einträgen, die solche Werkzeuge erst interessant machen. Neue Einträge ins Glossar kommen in der Regel dann hinzu, wenn

im Zug der Bearbeitung von Einheiten des Grammatischen Grundwissens bislang nicht aufgetretene Fachbegriffe gebraucht werden. Auch die Einträge ins Wörterbuch werden nicht losgelöst von den sonstigen Arbeiten vorgenommen, sondern immer dann, wenn eine einschlägige Komponente des Grundwissens fertig gestellt wird.

Zug um Zug werden weitere Einheiten fertig gestellt und online zugänglich gemacht. Es lohnt sich also, schon jetzt immer mal wieder bei GRAMMIS ›vorbeizubrowsen‹, um auf dem neuesten Stand zu bleiben. Anregungen und Kritik sind dem GRAMMIS-Team – Eva Breindl, Elke Donalies, Roman Schneider, Bruno Strecker (Projektleiter), Klaus Vorderwülbecke – jederzeit willkommen. Unsere Kontaktadresse: <http://www.ids-mannheim.de/grammis>.

Anmerkungen

¹ Zifonun, Gisela/Hoffmann, Ludger/Strecker, Bruno et al. (1997): Grammatik der deutschen Sprache. 3 Bände. Berlin/New York

² Für diese Möglichkeit sind wir Peter Eisenberg und seinen Mitarbeitern an der Universität Potsdam zu besonderem Dank verpflichtet, weil sie die von ihnen entwickelte Bibliographie mit ca. 10.000 Titeln dem Institut für Deutsche Sprache zur Weiterführung und Pflege überlassen haben.

³ Die ›Wissensbasis‹, auf der diese GRAMMIS-Einheiten aufbauen, ist, mit Ausnahme der Darstellung der Wortbildung, die Grammatik der deutschen Sprache, doch handelt es sich dabei um mehr als eine einfache Umsetzung des Druckwerks. Die Schichtung der Information in drei Ebenen, die erweiterten Darstellungsmöglichkeiten und nicht zuletzt die Bearbeitung des Gegenstands durch andere Autoren führt auch in der Sache zu Änderungen, die weiter gehen, als ursprünglich zu erwarten war.

Dr. Eva Breindl, Roman Schneider und Prof. Dr. Bruno Strecker sind wissenschaftliche Mitarbeiter im Institut für Deutsche Sprache in Mannheim.